APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: CROSS CORRELATED TRELLIS CODED QUADRATURE MODULATION TRANSMITTER AND SYSTEM

APPLICANT: MARVIN K. SIMON AND TSUN-YEE YAN

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL528178996US

I hereby certify that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Date of Deposit: February 1, 2000

Signature

Typed or Printed Name of Person Signing Certificate

CROSS CORRELATED TRELLIS CODED QUADRATURE MODULATION TRANSMITER AND SYSTEM

INS A'

A'

Cross Reference To Related Applications

This application claims the benefit of the U.S. Provisional Application No. 60/103,227, filed on October 5, 1998.

Background

Information can be sent over a channel using modulation techniques. Better bandwidth efficiency allows this same channel to hold and carry more information. A number of different systems for efficiently transmitting over channels are known. Examples include Gaussian minimum shift keying, staggered quadrature overlapped raised cosine modulation, and Feher's patented quadrature phase shift keying.

Many of these systems provide a transmitted signal with a constant or pseudo-constant envelope. This is desirable when the transmitter has a nonlinear amplifier that operates in or near saturation.

Many of these phase shift keying signals systems can operate using limited groups of the information at any one time.

Trellis coded modulation techniques are well known. Trellis coded techniques operate using multi-level modulation techniques, and hence can be more efficient than symbol-by-symbol transmission techniques.

Summary

The present application teaches a special cross correlated trellis coded quadrature modulation technique that can be used with a variety of different transmission schemes. Unlike conventional systems that use constant envelopes for the modulating waveforms, the present system enables mapping onto an arbitrarily chosen waveform that is selected based on bandwidth efficiency for the particular channel.

The system uses a special cross correlator that carries out the mapping in a special way.

This system can be used with offset quadrature phase shift keying along with conventional encoders, matched filters, decoders and the like. The system uses a special form of trellis coding in the modulation process that shapes

the power spectrum of the transmitted signal over and above bandwidth efficiency that is normally achieved by an M-ary (as opposed to binary) modulation.

Brief Description of the Drawings

These and other aspects of the invention will be described in detail with reference to the accompanying drawings, wherein:

Figure 1 shows a basic block diagram of a preferred transmitter of the present application;

Figure 2 shows a specific cross correlation mapper;

Figure 3 shows a specific embodiment that is optimized for XPSK;

Figure 4 shows waveforms for FQPSK;

Figure 5 shows a block diagram of the system for FQPSK;

Figure 6a and 6b respectively show the waveforms for in phase and out of phase FQPSK outputs;

Figure 7 shows a trellis diagram for FQPSK;

Figure 8 shows an FQPSK shaper;

Figure 9 shows waveforms for full symbols of OQPSK;

Figure 10 shows a trellis coded OQPSK;

Figure 11 shows a 2 state trellis diagram;

Figure 12 shows an uncoded OQPSK transmitter; and

Figure 13 shows paths.

Detailed Description

The present application describes a system with a transmitter that can operate using trellis coding techniques, which improve the operation as compared with the prior art techniques.

The present application focuses on the spectral occupancy of the transmitted signal. A special envelope property is described that improves the power efficiency of the demodulation and decoding operation. The disclosed structure is generic, and can be applied to different kinds of modulation including XPSK, FQPSK, SQORC, MSK and OP or OQPSK.

Fig. 1 shows a block diagram of a cross correlated quadrature modulation (XTCQM) transmitter 100.

An input binary (±1) datastream 105 is an independent, identically distributed information sequence $\{d_n\}$ at a bit rate $R_b = 1/T_b$. A quadrature converter 110 separates this sequence into an inphase (I) sequence 102 and a quadriphase (Q) sequence 104 $\{d_{In}\}$ and $\{d_{Qn}\}$. As conventional, every

second bit becomes part of the different phase. Hence, the phases can be formed by the even and odd bits of the information bit sequence $\{d_n\}$. The bits hence occur on the I and Q channels at a rate $R_s = 1/T_s = 1/2T_b$; where T_b is the bit rate, and T_s is the symbol rate.

For this explanation, it is assumed that the I and Q sequences $\{d_{In}\}$ and $\{d_{Qn}\}$ are time synchronous. Hence, each bit d_{In} (or d_{Qn}) occurs during the interval $(n-\frac{1}{2})T_s \le t \le (n+\frac{1}{2})T_s$ where n represents a count of adjacent symbol time periods T_s.

Rather than analyzing these levels as extending from +1 to -1, it may be more convenient to work with the (0,1) equivalents of the I and Q data sequences. This can be defined as

$$D_{In} \triangleq \frac{1-d_{In}}{2}, \qquad D_{Qn} \triangleq \frac{1-d_{Qn}}{2} \tag{1}$$

which both range within the set (0,1). The sequences $\{D_{In}\}$ and $\{D_{Qn}\}$ are separately and respectively applied to rate $r = 1/N$ convolutional encoders 120, 125. The two encoders are in general different, i.e., they have different tap connections and different modulo 2 summers but are assumed to have the same code rate.

We can define $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$, $\left\{E_{Qk}\middle|_{k=1}^{N}\right\}$ respectively as the sets of $N(0,1)$ output symbols 122, 127 respectively, of the I and Q convolutional encoders 120, 125 corresponding to a single bit input to each of the encoders.

These sets of output symbols 122, 127 will be used to determine a pair of baseband waveforms $s_I(t), s_Q(t)$ which ultimately modulate I and Q carriers for transmission over the channel. The signal $s_Q(t)$ is delayed by delay element 130 for $T_s/2 = T_b$ seconds prior to modulation on the quadrature carrier. This delay offsets the signal $s_Q(t)$ relative to the $s_I(t)$ signal, and thereby provides an offset modulation. Delaying the waveform by one half of a symbol at the output of the mapping allows synchronous demodulation and facilitates computation of the path metric at the receiver. This is different than the approach used for conventional FQPSK.

The present application teaches mapping of the symbol sets $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$ and $\left\{E_{Qk}\middle|_{k=1}^{N}\right\}$ into $s_I(t)$ and $s_Q(t)$ using a waveform with a desired size and content ("waveshape").

Mapping

The mapping of the sets $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$ and $\left\{E_{Qk}\middle|_{k=1}^{N}\right\}$ into $s_I(t)$ and $s_Q(t)$ uses a crosscorrelation mapper 140. Details of the mapping is shown in Fig. 2. Each of these sets of N (0,1) output symbols is partitioned into one of three groups as follows.

The I and Q signals are separately processed. For the I signals, the first group uses $I_{I_1}, I_{I_2}, \ldots, I_{I_{N_1}}$ as a subset of N_1 elements of $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$ which will be used only in the selection of $s_I(t)$. The second group uses $Q_{I_1}, Q_{I_2}, \ldots, Q_{I_{N_2}}$ as a subset N_2 elements of $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$ which will be used *only* in the selection of $s_Q(t)$. The third group uses $I_{I_{N_1+1}}, I_{I_{N_1+2}}, \ldots, I_{I_{N_1+N_3}} = Q_{I_{N_2+1}}, Q_{I_{N_2+2}}, \ldots, Q_{I_{N_2+N_3}}$ as a subset of N_3 elements of $\left\{E_{Ik}\middle|_{k=1}^{N}\right\}$ which will be used *both* for the selection of $s_I(t)$ and $s_Q(t)$. The term "crosscorrelation" in this context refers to the way in which the groups are formed.

All of the output symbols of the I encoder are used either to select $s_I(t)$, $s_Q(t)$ or both. Therefore,

$N_1 + N_2 + N_3 = N$.

A similar three part grouping of the Q encoder output symbols $\left\{ E_{Qk} \middle| \begin{matrix} N \\ k=1 \end{matrix} \right\}$ occurs. That is, for the first group let $Q_{m_1}, Q_{m_2}, \ldots, Q_{m_{L_1}}$ be a subset L_1 elements of $\left\{ E_{Qk} \middle| \begin{matrix} N \\ k=1 \end{matrix} \right\}$ which will be used *only* in the selection of $s_Q(t)$. For the second group, let $I_{m_1}, I_{m_2}, \ldots, I_{m_{L_2}}$ be a subset of L_2 elements of $\left\{ E_{Qk} \middle| \begin{matrix} N \\ k=1 \end{matrix} \right\}$ which will be used *only* in the selection of $s_I(t)$. Finally, for the third group let $Q_{m_{L_1+1}}, Q_{m_{L_1+2}}, \ldots, Q_{m_{L_1+L_3}} = I_{m_{L_2+1}}, I_{m_{L_2+2}}, \ldots, I_{m_{L_2+L_3}}$ be a subset of L_3 elements of $\left\{ E_{Qk} \middle| \begin{matrix} N \\ k=1 \end{matrix} \right\}$ which will be used *both* for the selection of $s_I(t)$ and $s_Q(t)$. Once again, since all of the output symbols of the Q encoder are used either to select $s_I(t)$, $s_Q(t)$ or both, then $L_1 + L_2 + L_3 = N$.

A preferred mode exploits symmetry properties associated with the resulting modulation by choosing $L_1 = N_1$, $L_2 = N_2$ and $L_3 = N_3$. However, the present invention is not restricted to this particular symmetry.

In summary, based on the above, the signal $S_k(t)$ is determined from symbols $I_{I_1}, I_{I_2}, \ldots, I_{I_{N_1+N_3}}$ from the output of the I

encoder and symbols $I_{l_1}, I_{l_2}, \ldots, I_{l_{L_2}+L_3}$ from the output of the Q encoder. Thus, the size of the signaling alphabet used to select $s_I(t)$ is $2^{N_1+N_3+L_2+L_3} \triangleq 2^{N_I}$. Similarly, the signal $s_Q(t)$ is determined from symbols $Q_{l_1}, Q_{l_2}, \ldots, Q_{l_{L_1}+L_3}$ from the output of the Q encoder and symbols $Q_{l_1}, Q_{l_2}, \ldots Q_{l_{N_2}+N_3}$ from the output of the I encoder. Thus, the size of the signaling alphabet used to select $s_Q(t)$ is $2^{L_1+L_3+N_2+N_3} \triangleq 2^{N_Q}$.

An interesting embodiment results when the size of the signaling alphabets for selecting $s_I(t)$ and $s_Q(t)$ are equal. In that case, $N_I = N_Q$ or equivalently $L_1 + N_2 = N_1 + L_2$. This condition is clearly satisfied if the condition $L_1 = N_1, L_2 = N_2$ is met; however, the former condition is less restrictive and does not require the latter to be true.

Fig. 3 shows an example of the above mapping corresponding to $N_1 = N_2 = N_3 = 1$ and $L_1 = L_2 = L_3 = 1$, i.e., $r = 1/N = 1/3$ encoders for FQPSK, which is one particular embodiment of the XTCQM invention. The specific symbol assignments for the three partitions of the I encoder output are I_3 (group 1), Q_0 (group 2), $I_2 = Q_1$ (group 3). Similarly, the specific symbol assignments for the three

partitions of the Q encoder output are: Q_3 (group 1), I_1 (group 2), $I_0 = Q_2$ (group 3). Since $N_I = N_Q = 4$, the size of the signaling alphabet from which both $s_I(t)$ and $s_Q(t)$ are to be selected has 2^4 = 16 signals.

After assigning the encoder output symbols to either $s_I(t)$, $s_Q(t)$ or both, appropriate binary coded decimal (BCD) numbers are formed from these symbols. These numbers are used as indices i and j for selecting $s_I(t) = s_i(t)$ and $s_Q(t) = s_j(t)$ where $\left\{ s_i(t) \middle| {}^{N_I}_{i=1} \right\}$ and $\left\{ s_j(t) \middle| {}^{N_Q}_{j=1} \right\}$ are the signal waveform sets assigned for transmission of the I and Q channel signals.

$I_0, I_1, \ldots, I_{N_I}$ are defined as the specific set of symbols taken from both I and Q encoder outputs used to select $s_I(t)$ and $s_Q(t)$. Then the BCD indices needed above are $i = I_{N_I - 1} \times 2^{N_I - 1} + \cdots + I_1 \times 2^1 + \cdots + I_0 \times 2^0$ and $j = Q_{N_Q - 1} \times 2^{N_Q - 1} + \cdots + Q_1 \times 2^1 + \cdots + Q_0 \times 2^0$. The Fig. 2 embodiment uses $i = I_3 \times 2^3 + I_2 \times 2^2 + I_1 \times 2^1 \cdots + I_0 \times 2^0$ and $j = Q_3 \times 2^3 + Q_2 \times 2^2 + Q_1 \times 2^1 \cdots + Q_0 \times 2^0$. This is shown in Figure 3.

Numerically speaking, in a particular transmission interval of T_s seconds, the contents of the I and Q encoders

in Fig. 3 can be $D_{I,n+1}=1, D_{In}=0, D_{I,n-1}=0$ and $D_{Q,n}=1, D_{Q,n-1}=0, D_{Q,n-2}=1$, then the encoder output symbols $\left\{E_{Ik}\middle|_{k=1}^{3}\right\}$ and $\left\{E_{Qk}\middle|_{k=1}^{3}\right\}$ would respectively partition as $I_3=0$ (group 1), $Q_0=1$ (group 2), $I_2=Q_1=0$ (group 3) and $Q_3=1$ (group 1), $I_1=1$ (group 2), $I_0=Q_2=1$ (group 3). Thus, based on the above, $i=3$ and $j=13$ and hence the selection for $s_I(t)$ and $s_Q(t)$ would be $s_I(t)=s_3(t)$ and $s_Q(t)=s_{13}(t)$.

The Signal Sets (Waveforms)

An important function of the present application is that any set of N_I waveforms of duration T_s seconds (defined on the interval $(-T_s/2 \le t \le T_s/2)$ can be used for selecting the I channel transmitted signal. Likewise, any set of N_Q waveforms of duration T_s seconds, also defined on the interval $(-T_s/2 \le t \le T_s/2)$ can be used for selecting the Q channel transmitted signal $s_Q(t)$. However, certain properties can be invoked on these waveforms to make them more power and spectrally efficient.

This discussion assumes the special case of $N_I = N_Q \triangleq N^*$, although other embodiments are contemplated. Maximum

distance in the waveform set can improve power efficiency. The distance can be increased by dividing the signal set $\left\{ s_i(t) \middle| \begin{matrix} N^* \\ i=1 \end{matrix} \right\}$ into two equal parts; with the signals in the second part being antipodal to (the negatives of) those in the first part. Mathematically, the signal set has the composition $s_0(t), s_1(t), \ldots, s_{N^*/2-1}(t), -s_0(t), -s_1(t), \ldots, -s_{N^*/2-1}(t)$. To achieve good spectral efficiency, one should choose the waveforms to be as smooth, i.e., as many continuous derivatives, as possible, since a smoother waveform gives better power spectrum roll off. Furthermore, to prevent discontinuities at the symbol transition time instants, the waveforms should have a zero first derivative (slope) at their endpoints $t = \pm T_s / 2$.

An example of a signal set that satisfies the first requirement and part of the second requirement is still illustrated in Fig. 4. This shows the specific FQPSK embodiment.

Conventional FQPSK

Generic FQPSK is described in U.S. Patent numbers 4,567,602; 4,339,724; 4,644,565 and 5,491,457. This is *conceptually* similar to the *cross-correlated phase-shift-keying* (XPSK) modulation technique introduced in 1983 by

Kato and Feher. This technique was in turn a modification of the previously-introduced (by Feher et al) *interference and jitter free* QPSK (IJF-QPSK) with the purpose of reducing the 3 dB envelope fluctuation characteristic of IJF-QPSK to 0 dB. This made the modulation appear as a constant envelope, which was beneficial in nonlinear radio systems. It is further noted that using a constant waveshape for the even pulse and a sinusoidal waveshape for the odd pulse, IJF-QPSK becomes identical to the staggered quadrature overlapped raised cosine (SQORC) scheme introduced by Austin and Chang. Kato and Feher achieved their 3 dB envelope reduction by using an intentional but controlled amount of crosscorrelation between the inphase (I) and quadrature (Q) channels. This crosscorrelation operation was applied to the IJF-QPSK (SQORC) baseband signal prior to its modulation onto the I and Q carriers.

Fig. 5 shows a conceptual block diagram of FPQSK. Specifically, this operation has been described by mapping, *in each half symbol,* the 16 possible combinations of I and Q channel waveforms present in the SQORC signal. The mapping moves the signals into a new set of 16 waveform combinations chosen in such a way that the crosscorrelator output is time

continuous and has a unit (normalized) envelope at all I and Q uniform sampling instants.

The present embodiment describes restructuring the crosscorrelation mapping into one mapping, based on a *full symbol* representation of the I and Q signals. The FQPSK signal can be described directly in terms of the data transitions on the I and Q channels. As such, the representation becomes a specific embodiment of XTCQM.

Appropriate mapping of the transitions in the I and Q data sequences into the signals $s_I(t)$ and $s_Q(t)$ is described by Tables 1 and 2.

Table 1. Mapping for Inphase (I)-Channel Baseband Signal $s_I(t)$ in the Interval $(n-\frac{1}{2})T_s \le t \le (n+\frac{1}{2})T_s$

$\left\|\frac{d_{In}-d_{I,n-1}}{2}\right\|$	$\left\|\frac{d_{Q,n-1}-d_{Q,n-2}}{2}\right\|$	$\left\|\frac{d_{Qn}-d_{Q,n-1}}{2}\right\|$	$s_I(t)$
0	0	0	$d_{In}s_0(t-nT_s)$
0	0	1	$d_{In}s_1(t-nT_s)$
0	1	0	$d_{In}s_2(t-nT_s)$
0	1	1	$d_{In}s_3(t-nT_s)$
1	0	0	$d_{In}s_4(t-nT_s)$
1	0	1	$d_{In}s_5(t-nT_s)$
1	1	0	$d_{In}s_6(t-nT_s)$
1	1	1	$d_{In}s_7(t-nT_s)$

Table 2. Mapping for Quadrature (Q)-Channel Baseband Signal $s_Q(t)$ in the Interval $(n-\frac{1}{2})T_s \le t \le (n+\frac{1}{2})T_s$

$\left\|\frac{d_{Qn}-d_{Q,n-1}}{2}\right\|$	$\left\|\frac{d_{I,n}-d_{I,n-1}}{2}\right\|$	$\left\|\frac{d_{I,n+1}-d_{I,n}}{2}\right\|$	$s_Q(t)$
0	0	0	$d_{Qn}s_0(t-nT_s)$
0	0	1	$d_{Qn}s_1(t-nT_s)$
0	1	0	$d_{Qn}s_2(t-nT_s)$
0	1	1	$d_{Qn}s_3(t-nT_s)$
1	0	0	$d_{Qn}s_4(t-nT_s)$
1	0	1	$d_{Qn}s_5(t-nT_s)$
1	1	0	$d_{Qn}s_6(t-nT_s)$
1	1	1	$d_{Qn}s_7(t-nT_s)$

Note that the subscript i of the transmitted signal $s_I(t)$ *or* $s_Q(t)$ *as appropriate is the binary coded decimal (BCD) equivalent of the three transitions.* Since d_{Im} and d_{Qm} take on values ±1, Tables 1 and 2 specify the mapping of I and Q symbol transitions 16 different waveforms, namely, $s_i(t)\big|_{i=0}^{15}$ where $s_i(t) = -s_{i-8}(t), i = 8,9,\ldots,15$.

The specifics are as follows:

$$s_0(t) = A, \quad -T_s/2 \le t \le T_s/2, \qquad s_8(t) = -s_0(t)$$

$$s_1(t) = \begin{cases} A, & -T_s/2 \le t \le 0 \\ 1-(1-A)\cos^2\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases} \qquad s_9(t) = -s_1(t) \tag{2a}$$

$$s_2(t) = \begin{cases} 1-(1-A)\cos^2\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ A, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{10}(t) = -s_2(t)$$

$$s_3(t) = 1-(1-A)\cos^2\frac{\pi t}{T_s}, -T_s/2 \le t \le T_s/2 \qquad s_{11}(t) = -s_3(t)$$

and

$$s_4(t) = A\sin\frac{\pi t}{T_s}, \quad -T_s/2 \le t \le T_s/2, \qquad s_{12}(t) = -s_4(t)$$

$$s_5(t) = \begin{cases} A\sin\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ \sin\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{13}(t) = -s_5(t)$$

$$s_6(t) = \begin{cases} \sin\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ A\sin\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{14}(t) = s_6(t) \tag{2b}$$

$$s_7(t) = \sin\frac{\pi t}{T_s}, \quad -T_s/2 \le t \le T_s/2, \qquad s_{15}(t) = s_7(t)$$

Applying the mappings in Tables 1 and 2 to the I and Q data sequences produces the identical I and Q baseband transmitted signals to those that would be produced by passing the I and Q IJF encoder outputs of Figure 5 through the crosscorrelator (half symbol mapping) of the FQPSK (XPSK) scheme. An example of this is shown with reference to Figures 6a and 6b. The Q signal must be delayed by $T_s/2$ to produce an offset form of modulation. Alternately stated, *for arbitrary I and Q data sequences, FQPSK can alternately be generated by the symbol-by-symbol mappings of Tables 1 and 2 as applied to these sequences.*

The mappings of Tables 1 and 2 become a specific embodiment of XTCQM as described herein. First, the I and Q transitions needed for the BCD representations of the

indices of $s_i(t)$ *and* $s_j(t)$ are rewritten in terms of their modulo 2 sum equivalents. That is, using the (0,1) form of the I and Q data symbols, Tables 1 and 2 show that

$$i = I_3 \times 2^3 + I_2 \times 2^2 + I_1 \times 2^1 + I_0 \times 2^0$$
$$j = Q_3 \times 2^3 + Q_2 \times 2^2 + Q_1 \times 2^1 + Q_0 \times 2^0 \qquad (3)$$

with

$$\begin{aligned} I_0 &= D_{Qn} \oplus D_{Q,n-1}, \quad Q_0 = D_{I,n+1} \oplus D_{In} \\ I_1 &= D_{Q,n-1} \oplus D_{Q,n-2}, \quad Q_1 = D_{In} \oplus D_{I,n-1} = I_2 \\ I_2 &= D_{In} \oplus D_{I,n-1}, \quad Q_2 = D_{Qn} \oplus D_{Q,n-1} = I_0 \\ I_3 &= D_{In}, \quad Q_3 = D_{Qn} \end{aligned} \qquad (4)$$

resulting in the baseband I and Q waveforms $s_I(t) = s_i(t - nT_s)$ and $s_Q(t) = s_j(t - nT_s)$. The signals that are modulated onto the I and Q carriers are $y_I(t) = s_I(t)$ and $y_Q(t) = s_Q(t - T_s/2)$. Thus, in each symbol interval $((n - \frac{1}{2})T_s \le t \le (n + \frac{1}{2})T_s$ for $y_I(t)$ and $nT_s \le t \le (n+1)T_s$ for $y_Q(t)$), the I and Q channel baseband signals are each chosen from a set of 16 signals, $s_i(t), i = 0,1,\ldots,15$ in accordance with the 4-bit BCD representations of their indices defined by (3) together with (4).

A graphical illustration of the implementation of this mapping is given in Figure 3, which is a specific embodiment

of Figure 1 with $N_1 = N_2 = N_3 = L_1 = L_2 = L_3 = 1$. The mapping in Figure 3 can be interpreted as a 16-state trellis code with two binary inputs $D_{I,n+1}, D_{Qn}$ and two waveform outputs $s_i(t), s_j(t)$ where the state is defined by the 4-bit sequence $D_{In}, D_{I,n-1}, D_{Q,n-1}, D_{Q,n-2}$. The trellis is illustrated in Figure 7 and the transition mapping is given in Table 3.

Table 3. Trellis State Transistions

Current State	Input	Output	Next State
0 0 0 0	0 0	0 0	0 0 0 0
0 0 0 0	0 1	1 12	0 0 1 0
0 0 0 0	1 0	0 1	1 0 0 0
0 0 0 0	1 1	1 13	1 0 1 0
0 0 1 0	0 0	3 4	0 0 0 1
0 0 1 0	0 1	2 8	0 0 1 1
0 0 1 0	1 0	3 5	1 0 0 1
0 0 1 0	1 1	2 9	1 0 1 1
1 0 0 0	0 0	12 3	0 1 0 0
1 0 0 0	0 1	13 15	0 1 1 0
1 0 0 0	1 0	12 2	1 1 0 0
1 0 0 0	1 1	13 14	1 1 1 0
1 0 1 0	0 0	15 7	0 1 0 1
1 0 1 0	0 1	14 11	0 1 1 1
1 0 1 0	1 0	15 6	1 1 0 1
1 0 1 0	1 1	14 10	1 1 1 1
0 0 0 1	0 0	2 0	0 0 0 0
0 0 0 1	0 1	3 12	0 0 1 0
0 0 0 1	1 0	2 1	1 0 0 0
0 0 0 1	1 1	3 13	1 0 1 0
0 0 1 1	0 0	1 4	0 0 0 1
0 0 1 1	0 1	0 8	0 0 1 1
0 0 1 1	1 0	1 5	1 0 0 1
0 0 1 1	1 1	0 9	1 0 1 1
1 0 0 1	0 0	14 3	0 1 0 0
1 0 0 1	0 1	15 15	0 1 1 0
1 0 0 1	1 0	14 2	1 1 0 0
1 0 0 1	1 1	15 14	1 1 1 0
1 0 1 1	0 0	13 7	0 1 0 1
1 0 1 1	0 1	12 11	0 1 1 1
1 0 1 1	1 0	13 6	1 1 0 1
1 0 1 1	1 1	12 10	1 1 1 1
0 1 0 0	0 0	4 2	0 0 0 0
0 1 0 0	0 1	5 14	0 0 1 0
0 1 0 0	1 0	4 3	1 0 0 0
0 1 0 0	1 1	5 15	1 0 1 0
0 1 1 0	0 0	7 6	0 0 0 1
0 1 1 0	0 1	6 10	0 0 1 1
0 1 1 0	1 0	7 7	1 0 0 1
0 1 1 0	1 1	6 11	1 0 1 1
1 1 0 0	0 0	8 1	0 1 0 0
1 1 0 0	0 1	9 13	0 1 1 0
1 1 0 0	1 0	8 0	1 1 0 0
1 1 0 0	1 1	9 12	1 1 1 0

	1	1	1	0		0	0		11	5		0	1	0	1
	1	1	1	0		0	1		10	9		0	1	1	1
	1	1	1	0		1	0		11	4		1	1	0	1
	1	1	1	0		1	1		10	8		1	1	1	1
5	0	1	0	1		0	0		6	2		0	0	0	0
	0	1	0	1		0	1		7	14		0	0	1	0
	0	1	0	1		1	0		6	3		1	0	0	0
	0	1	0	1		1	1		7	15		1	0	1	0
	0	1	1	1		0	0		5	6		0	0	0	1
10	0	1	1	1		0	1		4	10		0	0	1	1
	0	1	1	1		1	0		5	7		1	0	0	1
	0	1	1	1		1	1		4	11		1	0	1	1
	1	1	0	1		0	0		10	1		0	1	0	0
	1	1	0	1		0	1		11	13		0	1	1	0
15	1	1	0	1		1	0		10	0		1	1	0	0
	1	1	0	1		1	1		11	12		1	1	1	0
	1	1	1	1		0	0		9	5		0	1	0	1
	1	1	1	1		0	1		8	9		0	1	1	1
	1	1	1	1		1	0		9	4		1	1	0	1
20	1	1	1	1		1	1		8	8		1	1	1	1

In this table, the entries in the column labeled "input" correspond to the values of the two input bits $D_{I,n+1}, D_{Qn}$ that result in the transition. The entries in the column "output" correspond to the subscripts *i* and *j* of the pair of symbol waveforms $s_i(t), s_j(t)$ that are output.

Enhanced FQPSK

It is well known that the rate at which the sidelobes of a modulation's power spectral density (PSD) roll off with frequency is related to the smoothness of the underlying waveforms that generate it. That is, a waveform that has more continuous waveform derivatives will have faster Fourier transform decays with frequency.

The crosscorrelation mappings of FQPSK is based on a half symbol characterization of the SQORC signal. Hence, there is no guarantee that the *slope* or any higher derivatives of the crosscorrelator output waveform is continuous at the half symbol transition points. From Equation (2b) and the corresponding illustration in Figure 4, it can be observed that four out of the sixteen possible transmitted waveforms, namely, $s_5(t), s_6(t), s_{13}(t), s_{14}(t)$ have a slope discontinuity at their midpoint. Thus, for random I and Q

data symbol sequences, on the average the transmitted FQPSK waveform will likewise have a slope discontinuity at one quarter of the uniform sampling time instants. Therefore, for a random data input sequence, a discontinuity in slope occurs one quarter of the time.

Based on the above reasoning, it is predicted that an improvement in PSD rolloff could be obtained if the FQPSK crosscorrelation mapping could be modified so that the first derivative of the transmitted baseband waveforms is always continuous. This enhanced version of FQPSK requires a slight modification of the above-mentioned four waveforms in Figure 4. In particular, these four transmitted signals are redefined in a manner analogous to $s_1(t), s_2(t), s_9(t), s_{10}(t)$, namely

$$s_5(t) = \begin{cases} \sin\dfrac{\pi t}{T_s} + (1-A)\sin^2\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ \sin\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{13}(t) = -s_5(t)$$

$$s_6(t) = \begin{cases} \sin\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ \sin\dfrac{\pi t}{T_s} - (1-A)\sin^2\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{14}(t) = -s_6(t) \qquad (5)$$

Note that not only do the signals $s_5(t), s_6(t), s_{13}(t), s_{14}(t)$ as defined in (5) *not* have a slope discontinuity at their midpoint, or anywhere else in the defining interval. Also, the zero slope at their endpoints has been preserved. Thus,

the signals in (5) satisfy both requirements for desired signal set waveforms as discussed in Section 3.1.2. Using (5) in place of the corresponding signals of (2b) results in a modified FQPSK signal that has no slope discontinuity anywhere in time *regardless of the value of A*.

Figure 5 illustrates a comparison of the signal $s_0(t)$ of (5) with that of (2b) for a value of $A = 1/\sqrt{2}$.

The signal set selected for enhanced FQPSK has a symmetry property for $s_0(t) - s_3(t)$ that is not present for $s_4(t) - s_7(t)$. In particular, $s_1(t)$ and $s_2(t)$ are each composed of one half of $s_0(t)$ and one half of $s_3(t)$, i.e., the portion of $s_1(t)$ from $t = -T_s/2$ to $t = 0$ is the same as that one half of $s_0(t)$ whereas the portion of $s_1(t)$ from $t = 0$ to $t = T_s/2$ is the same as that of $s_3(t)$ and vice versa for $s_2(t)$. To achieve the same symmetry property for $s_4(t) - s_7(t)$, one would have to reassign $s_4(t)$ as

$$s_4(t) = \begin{cases} \sin\dfrac{\pi t}{T_s} + (1-A)\sin^2\dfrac{\pi t}{T_s}, & -T_s/2 \le t \le 0 \\ \sin\dfrac{\pi t}{T_s} - (1-A)\sin^2\dfrac{\pi t}{T_s}, & 0 \le t \le T_s/2 \end{cases}, \qquad s_{12}(t) = -s_4(t) \tag{6}$$

This minor change produces a complete symmetry in the waveform set. Thus, it has an advantage from the standpoint of hardware implementation and produces a negligible change in

spectral properties of the transmitted waveform. The remainder of the discussion, however, ignores this minor change and assumes the version of enhanced FQPSK first introduced in this section.

Trellis Coded OQPSK

Consider an XTCQM scheme in which the mapping function is performed identically to that in the FQPSK embodiment (i.e., as in Figure 3) but the waveform assignment is made as follows and as shown in Figure 9:

$$s_0(t) = s_1(t) = s_2(t) = s_3(t) = 1, \quad -T_s/2 \le t \le T_s/2,$$

$$s_4(t) = s_5(t) = s_6(t) = s_7(t) = \begin{cases} -1, & -T_s/2 \le t \le 0 \\ 1, & 0 \le t \le T_s/2 \end{cases} \tag{7}$$

$$s_i(t) = -s_{i-8}(t), i = 8, 9, \ldots, 15$$

that is, the first four waveforms are identical (a rectangular pulse) as are the second four (a split rectangular unit pulse) and the remaining eight waveforms are the negatives of the first eight. As such there are only four unique waveforms which are denoted by $c_i(t)\big|_{i=0}^{3}$ where $c_0(t) = s_0(t), c_1(t) = s_4(t), c_2(t) = s_8(t), c_3(t) = s_{12}(t)$. Since the BCD representations for each group of four identical waveforms

the two least significant bits are irrelevant, i.e., the two most significant bits are sufficient to define the common waveform for each group, the mapping scheme can be simplified by eliminating the need for I_0, I_1 and Q_0, Q_1. Fig. 3 shows how eliminating all of I_0, I_1 and Q_0, Q_1 accomplishes multiple purposes. The two encoders can be identical and need only a single shift register stage. Also, the correlation between the two encoders in so far as the mapping of either one's output symbols to both $s_I(t)$ and $s_Q(t)$ has been eliminated which therefore results in what might be termed a "degenerate" form of XTCQM.

The resulting embodiment is illustrated in Fig. 10. Since the mapping decouples the I and Q as indicated by the dashed line in the signal mapping block of Fig. 10, it is sufficient to examine the trellis structure and its distance properties for only one of the two I and Q channels. The trellis diagram for either channel of this modulation scheme would have two states as illustrated in Fig. 11. The dashed line indicates a transition caused by an input "0" and the solid indicates a transition caused by an input "1". Also, the branches are labeled with the output signal waveform that results from the transition. An identical trellis

diagram exists for the Q channel.

This embodiment of XTCQM has a PSD identical to that of the uncoded OQPSK (which is the same as uncoded QPSK) for the transmitted signal. In particular, because of the constraints imposed by the signal mapping, the waveforms $c_1(t) = s_4(t)$ and $c_3(t) = s_{12}(t)$ can never occur twice in succession. Thus, for any input information sequence, the sequence of signals $s_I(t)$ and $s_Q(t)$ cannot transition at a rate faster than $1/T_s$ sec. This additional spectrum conservation constraint imposed by the signal mapping function of XTCQM can reduce the coding (power) gain relative to that which could be achieved with another mapping which does not prevent the successive repetition of $c_1(t)$ and $c_3(t)$. However, the latter occurrence would result in a bandwidth expansion by a factor of two.

Trellis Coded SQORC

If instead of a split rectangular pulse in (7), a sinusoidal pulse were used, namely,

$$s_4(t) = s_5(t) = s_6(t) = s_7(t) = \sin\frac{\pi t}{T_s}, -T_s/2 \le t \le T_s/2$$

$$s_i(t) = -s_{i-8}(t), i = 12,13,14,15 \qquad (8)$$

then the same simplification of the mapping function as in Figure 10 occurs resulting in decoupling of the I and Q channels. The trellis diagram of Fig. 11 can then be used for either the I or Q channel. Once again, this has a PSD identical to that of uncoded SQORC which is the same as uncoded QORC.

Uncoded OQPSK

The signal assignment and mapping of Fig. 3 can be simplified such that

$$s_0(t) = s_1(t) = \cdots = s_7(t) = 1, \quad -T_s/2 \le t \le T_s/2,$$

$$s_i(t) = -s_{i-8}(t), \; i = 8,9,\ldots,15 \tag{9}$$

then in the BCD representations for each group of eight identical waveforms the three least significant bits are irrelevant. Only the first significant bit is needed to define the common waveform for each group. Hence, the mapping scheme can be simplified by eliminating the need for I_0, I_1, I_2 and Q_0, Q_1, Q_2. Defining the two unique waveforms $c_0(t) = s_0(t), c_1(t) = s_8(t)$ obtains the simplified degenerate mapping of Fig. 12 which corresponds to uncoded OQPSK with NRZ data formatting.

Likewise, if instead of the signal assignment in (9) the relation below is used:

$$s_0(t)+s_1(t)=\cdots=s_7(t)\begin{cases}-1, & -T_s/2\le t\le 0\\ 1, & 0\le t\le T_s/2\end{cases}$$
$$s_i(t)=-s_{i-8}(t),\ i=8,9,\ldots,15 \qquad (10)$$

then the mapping of Fig. 12 produces uncoded OQPSK with Manchester (biphase) data formatting.

Receiver Implementation and Performance

An optimum detector for XTCQM is a standard trellis coded receiver which employs a bank of filters which are matched to the signal waveform set, followed by a Viterbi (trellis) decoder. The bit error probability (BEP) performation of such a receiver can be described in terms of its minimum squared Euclidean distance $d^2_{\min}$, taken over all pairs of paths through the trellis. Comparing $d^2_{\min}$ for one TCM scheme with that of another scheme or with an uncoded modulation provides a measure of the relative *asymptotic coding gain in the limit of infinite* E_b/N_0. To compute $d^2_{\min}$ for a given TCM (of which XTCQM is one), it is sufficient to determine the minimum Euclidean distance over all pairs of error event paths that emanate from a given state, and first return to that or another state a number of branches later.

The procedure and actual coding gains that can be achieved relative to uncoded OQPSK are explained with reference to results for the specific embodiments of XTCQM discussed above.

FQPSK

For conventional or enhanced FQPSK, the smallest length error event for which there are at least two paths that start in one state and remerge in the same or another state is 3 branches. For each of the 16 starting states, there are exactly 4 such error event paths that remerge in each of the 16 end states. Fig. 13 is an example of these error event paths for the case where the originating state is "0000" and the terminating state is "0010".

The trellis code defined by the mapping in Table 3 is not uniform, e.g., it is not sufficient to consider only the all zeros path as the transmitted path in computing the minimum Euclidean distance. Rather all possible pairs of error event paths starting from each of the 16 states (the first 8 states are sufficient in view of the symmetry of the signal set) and the ending in each of the 16 states and must

be considered to determine the pair having the minimum Euclidean distance.

Upon examination of the squared Euclidean distance between all pairs of paths, regardless of length, it has been shown that the minimum of this distance normalized by the average bit energy which is one half the average energy of the signal (symbol) set, is for FQPSK given by

$$\frac{d_{\min}^2}{2\bar{E}_b} = \frac{16\left[\frac{7}{4} - \frac{8}{3\pi} - A\left(\frac{3}{2} + \frac{4}{3\pi}\right) + A^2\left(\frac{11}{4} + \frac{4}{\pi}\right)\right]}{\left(7 + 2A + 15A^2\right)} = 1.56 \qquad (11)$$

where $\bar{E}_b$ denotes the average bit energy of the FQPSK signal set, i.e., one-half the average symbol energy of the same signal set. For enhanced FQPSK we have

$$\frac{d_{\min}^2}{2\bar{E}_b} = \frac{(3 - 6A + 15A^2)}{\frac{21}{8} - \frac{8}{3\pi} - A\left(\frac{1}{4} - \frac{8}{3\pi}\right) + \frac{29}{8}A^2} = 1.56 \qquad (12)$$

which coincidentally is identical to that for FQPSK. Thus, the enhancement of FQPSK provided by using the waveforms of (5) as replacements for their equivalents in (2b) is significantly beneficial from a spectral standpoint with no penalty in asymptotic receiver performance.

To compare the performance of the optimum receivers of FQPSK and enhanced FQPSK with that of conventional uncoded offset QPSK (OQPSK) we note for the latter that $d^2_{\min}/\overline{E}_b = 2$ which is the same as that for BPSK. Thus, as a trade against the significantly improved power spectrum afforded by FQPSK and its enhanced version relative to that of OQPSK, *an asymptotic loss of only* 10 log(1/1.56)=1.07 *dB is experienced.* These results should be compared with the significantly poorer performance of the conventional FQPSK receiver which makes symbol-by-symbol decisions based independently on the I and Q samples, and results in an asymptotic loss in E_b/N_0 performance on the order of 2 to 2.5 dB relative to uncoded OQPSK.

Trellis Coded OQPSK

For the 2-state trellis diagram in Fig. 11, the minimum squared Euclidean distance occurs for an error event path of length 2 branches. Considering the four possible pairs of such paths that eminate from one of the 2 states and remerge at the same or the other state, then for the waveforms of Fig. 9 it is simple to see that $d^2_{\min} = 4T_s$. Since the average energy of the signal (symbol) set on the I (or Q) channel is

$E_{av} = T_s$ which is also equal to the average bit energy (since the channel by itself represents only one bit of information), then the normalized minimum squared Euclidean distance is $d^2_{\min} / 2\,\overline{E}_b = 2$ which represents *no asymptotic coding gain over OQPSK*. At finite values of E_b / N_0 there will exist some coding gain since the computation of error probability performance takes into account all possible error event paths, i.e., not only those corresponding to the minimum distance. Thus, in conclusion, the trellis coded OQPSK scheme presented here is a method for generating a transmitted modulation with a PSD that is identical to that of uncoded OQPSK and offers the potential of coding gain at finite SNR without the need for transmitting a higher order modulation (e.g., conventional rate 2/3 trellis coded 8PSK with also achieves no bandwidth expansion relative to uncoded QPSK), the latter being significant in that receiver synchronization circuitry can be designed for a quadriphase modulation scheme.

Trellis Coded SQORC

Here again the minimum squared Euclidean distance occurs for the same error event paths as described above.

With reference to the signal waveform, we now have $d^2_{\min} = 3T_s$. Since the average energy of this signal (symbol) set is $E_{av} = 0.75T_s$ which again per channel is equal to the average bit energy, then the normalized minimum squared Euclidean distance is also $d^2_{\min} / 2\,\overline{E}_b = 2$ which again represents *no asymptotic coding gain over SQORC*. Even though its pulse shaping SQORC has an improved PSD relative to OQPSK, it suffers from a 3 dB envelope fluctuation whereas OQPSK is constant envelope.